

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-306866

(43)Date of publication of application : 21.11.1995

(51)Int.Cl.

G06F 17/30  
G11B 27/00  
H04N 5/76

(21)Application number : 06-097441

(71)Applicant : NIPPON TELEGR &amp; TELEPH CORP &lt;NTT&gt;

(22)Date of filing : 11.05.1994

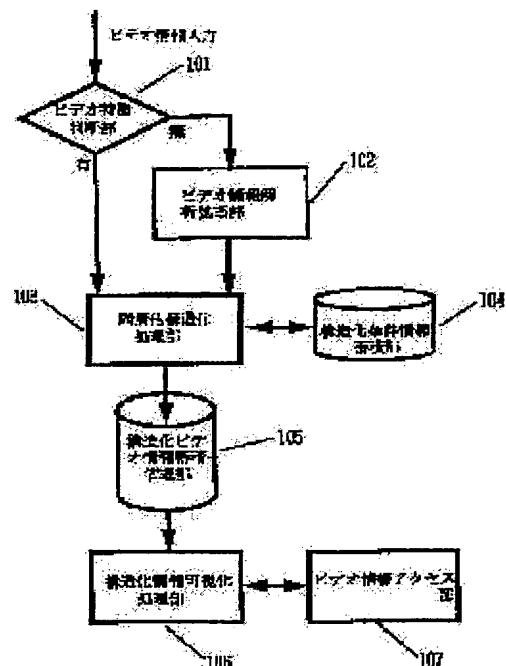
(72)Inventor : AKUTSU AKITO  
TONOMURA YOSHINOBU

## (54) VIDEO INFORMATION RETRIEVING METHOD

## (57)Abstract:

**PURPOSE:** To provide the video information retrieving method which performs a retrieving process instinctively and efficiently.

**CONSTITUTION:** This method is a method for retrieving specific video information in a video information source (102), structures the analysis result hierarchically (103), and visualizes the hierarchical structure of the video information source (106), thereby accessing the visualized video information. A video information analytic process 102 analyzes video images, speeches, subtitles, time codes, and an external input signal and extracts spatial variation features and time variation features of video information. A video structuring process 103 analyzes composes hierarchical video structure under structure conditions based upon the conditions wherein the images, speeches, subtitles, time codes, and external input signal of videos are analyzed and the spatial variation features and time variation features of the video information are extracted.



(19) 日本国特許庁 ( J P )

(12) 公 開 特 許 公 報 ( A )

(11) 特許出願公開番号

特開平7-306866

(43) 公開日 平成 7 年 ( 1995 ) 11 月 21 日

(51) Int.Cl. <sup>6</sup>	識別記号	序内整理番号	F I	技術表示箇所
G 0 6 F 17/30				
G 1 1 B 27/00	E	8224-5D		
H 0 4 N 5/76	B			
		9194-5L	G 0 6 F 15/ 40	3 7 0 D
		8224-5D	G 1 1 B 27/ 00	E
			審査請求 未請求	請求項の数 5 O L (全 6 頁)

(21) 出願番号 特願平6-97441

(22) 出願日 平成 6 年 ( 1994 ) 5 月 11 日

(71) 出願人 000004226

日本電信電話株式会社

東京都千代田区内幸町一丁目 1 番 6 号

(72) 発明者 阿久津 明人

東京都千代田区内幸町一丁目 1 番 6 号 日

本電信電話株式会社内

(72) 発明者 外村 佳伸

東京都千代田区内幸町一丁目 1 番 6 号 日

本電信電話株式会社内

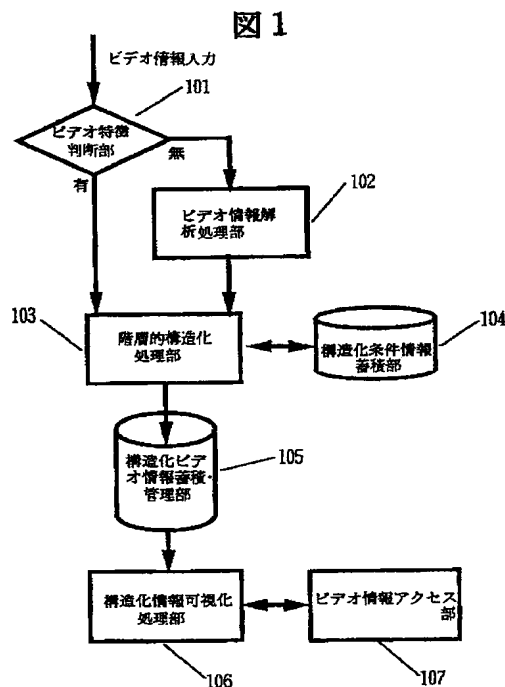
(74) 代理人 弁理士 秋田 収喜

(54) 【発明の名称】 ビデオ情報検索方法

(57) 【要約】

【目的】 検索過程を直感的にしかも効率よく行うことの可能なビデオ情報検索方法を提供する。

【構成】 ビデオの情報源から所定のビデオ情報を検索する方法であって、該ビデオ情報源のビデオ特徴を解析し、該解析結果を階層的に構造化し、該ビデオ情報源の階層構造を可視化し、該可視化ビデオ情報へアクセスする。前記ビデオ情報解析処理は、ビデオの画像、音声、字幕、タイムコード、外部入力信号を解析し、ビデオ情報の空間的变化特徴及び時間的变化特徴を抽出する。前記ビデオ構造化処理は、ビデオの画像、音声、字幕、タイムコード、外部入力信号を解析し、ビデオ情報の空間的变化特徴及び時間的变化特徴を抽出した条件による構造化条件を基に階層的に構造化する。



## 【特許請求の範囲】

【請求項 1】 ビデオの情報源から所定のビデオ情報を検索する方法であって、該ビデオ情報源のビデオ特徴を解析し、該解析結果を階層的に構造化し、該ビデオ情報源の階層構造を可視化し、該可視化ビデオ情報へアクセスすることを特徴とするビデオ情報検索方法。

【請求項 2】 請求項 1 に記載のビデオ情報検索方法において、前記ビデオ情報解析処理は、ビデオの画像、音声、字幕、タイムコード、外部入力信号を解析し、ビデオ情報の空間的変化特徴及び時間的変化特徴を抽出することを特徴とするビデオ情報検索方法。

【請求項 3】 請求項 1 または 2 に記載のビデオ情報検索方法において、前記ビデオ構造化処理は、ビデオの画像、音声、字幕、タイムコード、外部入力信号を解析し、ビデオ情報の空間的変化特徴及び時間的変化特徴を抽出した条件による構造条件を基に階層的に構造化することを特徴とするビデオ情報検索方法。

【請求項 4】 請求項 1 乃至 3 のうちいずれか 1 項に記載のビデオ情報検索方法において、前記ビデオ構造化処理は、ビデオの画像、音声、字幕、タイムコード、外部入力信号を解析し、ビデオ情報の空間的変化特徴を時空間解像度を基に階層的に構造化することを特徴とするビデオ情報検索方法。

【請求項 5】 請求項 1 乃至 4 のうちいずれか 1 項に記載のビデオ情報検索方法において、前記構造化情報可視化処理は、前記アクセスから入力された情報に対応し該ビデオ情報源の構造の可視化部分を自動的に随時変化させることを特徴とするビデオ情報検索方法。

## 【発明の詳細な説明】

## 【0001】

【産業上の利用分野】 本発明は、情報処理装置を用いてビデオ情報をインタラクティブに検索する場合のビデオ情報検索方法に関する。

## 【0002】

【従来の技術】 高度の情報化社会において、日々の活動はネットワークを通じ、情報処理装置（コンピュータ）を用いて情報の生成、検索、授受等を行いコミュニケーションをする。このようなコミュニケーション環境下において、情報の扱い易さがコミュニケーションの円滑化に大きく寄与することとなる。特に、氾濫した情報から利用者の所望の情報を容易に検索することの必要性が生じる。情報の検索に関してその形態を大きく 3 つに分類することができる。検索対象が明確であり、概念的に的確なキーワード等を用いて表現できる場合のキーワード検索をハンティングと呼ぶ。また、検索の目的意識無く試行錯誤的に所望の情報に近づく検索方法をブラウジングと呼ぶ。そして、ブラウジング同様検索の目的意識はなく、むしろあれこれ眺めながら発見的に情報を収集するタイプの検索をグレージングと呼ぶ。

【0003】 従来、情報がテキスト等の文書または書籍

の場合、情報は構造化することが技術的に可能であり、キーワードによるハンティング検索方法が広く用いられている。

【0004】 また、ブラウジングまたはグレージング検索の研究では、情報を仮想空間に三次元的遠近法を用いて表現し、遠く離れた所から情報を観察すると粗い大まかな情報が得られ、近づくにつれて細かく詳しい情報が得られるようなシステムが報告されている（Jock D. Mac kinlay, George G. Robertson and Stuart K. Card, The Perspective Wall: Detail and Context Smoothly Integrated, Proceedings of CHI'91 Human Factors in Computing Systems, pp.173-179, 1991、参照）。

【0005】 情報量の多いビデオを対象とした研究では、シーンチェンジを用いブラウジングする方法（大辻、外村、大庭：「輝度情報を使った動画ブラウジング」、信学技報、IE90-103, 1991、参照）やビデオを細かくアイコンメタファーで一覧表現し検索の容易性を実現している（外村、安部：「動画像データベースハンドリングに関する検討」、信学技報、IE89-33, 1989、参照）。

【0006】 また、検索したいビデオシーンのビデオ中の時間的な位置（場所）を検索者に意識させつつ、所望のシーンを検索する方法（Magnifier Tool）も報告されている（Michael Mills, Jonathan Cohen and Yin Yin Wong, A Magnifier Tool for Video Data, Proceedings of CHI'92 Human Factors in Computing Systems, pp.93-98, 1992、参照）。

## 【0007】

【発明が解決しようとする課題】 しかしながら、前記報告されている従来の技術または実現されているビデオ情報の検索方法（ブラウジング、グレージング）では、検索過程や所望のビデオシーンのビデオ中の時間的な位置（場所）及びビデオ構造上の位置（場所）を検索者に意識させつつ、効率よく検索することは不可能である。すなわち、仮想空間に三次元的遠近法を用い表現する報告では、ビデオに関して検討させておらず、シーンチェンジを用いブラウジングの報告では、検索過程や所望のビデオシーンをビデオ中の時間的な位置（場所）及びビデオ構造上の位置（場所）を検索者に意識させることは不可能であり、ビデオを細かくアイコンメタファーで一覧表現し検索する方法では、同様に検索過程や位置（場所）を検索者に意識させることは不可能であり、Magnifier Toolでは、ビデオの構造を積極的に用いていない。前述したように、従来報告されている技術は種々の問題点を抱えている。

【0008】 本発明は、前記問題点を解決するためになされたものであり、本発明の目的は、検索過程を直感的にしかも効率よく行うことが可能なビデオ情報検索方法を提供することにある。

【0009】 本発明の前記ならびにその他の目的及び新

規な特徴は、本明細書の記述及び添付図面によって明らかにする。

#### 【0010】

【課題を解決するための手段】本願において開示される発明のうち代表的なものの概要を簡単に説明すれば、以下のとおりである。

【0011】(1) ビデオの情報源から所定のビデオ情報を検索する方法であって、該ビデオ情報源のビデオ特徴を解析し、該解析結果を階層的に構造化し、該ビデオ情報源の階層構造を可視化し、該可視化ビデオ情報へアクセスすることを特徴とする。

【0012】(2) 前記(1)のビデオ情報検索方法において、前記ビデオ情報解析処理は、ビデオの画像、音声、字幕、タイムコード、外部入力信号を解析し、ビデオ情報の空間的変化特徴及び時間的変化特徴を抽出することを特徴とする。

【0013】(3) 前記(1)または(2)のビデオ情報検索方法において、前記ビデオ構造化処理は、ビデオの画像、音声、字幕、タイムコード、外部入力信号を解析し、ビデオ情報の空間的変化特徴及び時間的変化特徴を抽出した条件による構造条件を基に階層的に構造化することを特徴とする。

【0014】(4) 前記(1)乃至3のうちいずれか1つのビデオ情報検索方法において、前記ビデオ構造化処理は、ビデオの画像、音声、字幕、タイムコード、外部入力信号を解析し、ビデオ情報の空間的変化特徴を時空間解像度を基に階層的に構造化することを特徴とする。

【0015】(5) 前記(1)乃至(4)のうちいずれか1つの情報検索方法において、前記構造化情報可視化処理は、前記アクセス処理過程から入力された情報に対応し該ビデオ情報源の構造の可視化部分を自動的に随時変化させることを特徴とする。

#### 【0016】

【作用】前述の手段によれば、ビデオ情報源のビデオ特徴を解析し、該解析結果を階層的に構造化し、該ビデオ情報源の階層構造を可視化し、該可視化ビデオ情報へアクセスするビデオ情報検索方法において、前記ビデオ情報解析処理は、ビデオの画像、音声、字幕、タイムコード、外部入力信号を解析し、ビデオ情報の空間的変化特徴及び時間的変化特徴を抽出し、前記ビデオ構造化処理は、ビデオの画像、音声、字幕、タイムコード、外部入力信号を解析し、ビデオ情報の空間的変化特徴及び時間的変化特徴を抽出した条件による構造条件を基に階層的に構造化することより、ビデオ情報から所望のビデオまたはシーンを検索する場合に、検索者と所望情報との位置関係を検索者に意識させつつビデオ構造の積極的利用を行うので、検索過程を直感的にしかも効率よく行うことができる。

#### 【0017】

【実施例】以下、本発明の実施例を図面を参照して詳細

に説明する。

【0018】図1は本発明にかかる情報処理装置(コンピュータ)を用いてビデオ情報検索方法を行う一実施例の機能構成及びその処理手順を示すブロック機能構成図であり、101はビデオの特徴(インデクス)がビデオに付加されているかどうか判断するビデオ特徴判断部、102は特徴(インデクス)を抽出してビデオに付加するビデオ情報解析処理部、103は階層的に構造化する階層的構造化処理部、104は構造化条件情報蓄積部、105は構造化ビデオ情報蓄積・管理部、106は構造化ビデオ情報(階層構造)に基づいてビデオの可視化を行う構造化情報可視化処理部、107はビデオ情報アクセス部である。

【0019】本実施例のビデオ情報検索方法は、図1に示すように、入力されたビデオは、ビデオ特徴判断部101において、ビデオの特徴(インデクス)がビデオ情報に付加されているかどうかを判断する。もし、付加されていないときは、ビデオ情報解析処理部102において、前記ビデオ情報に入力ビデオの特徴(インデクス)を抽出して付加する。

【0020】ここで、ビデオ情報解析処理部102で抽出される特徴(インデクス)とは、ビデオ情報の物理的な特徴(輝度、色の変化、被写体の動き等)からビデオ構造に寄与する意味のある特徴(シーンチェンジ点、カメラ操作、ショット、シーン、ストーリー等)が自動または半自動で抽出され、記述されて入力ビデオに付加される。例えば、シーンチェンジ検出に関する方法として大辻他が報告している(K.Otsuji, Y.Tonomura, Projection Detecting Filter for Video Cut Detection, ACM Multimedia'93 Conference Proceedings, pp.251-257, 1993、参照)。

【0021】また、カメラ操作抽出に関して阿久津他が報告している(阿久津、外村、時空間画像を用いたグローバルな動き情報の抽出、1992年電子情報通信学会秋季大会、参照)。

【0022】被写体の認識によるビデオのインデクシングに関して上田他が報告している(上田他、動画像解析結果に基づく映像情報のブラウジング、1994年電子情報通信学会春季大会、参照)。

【0023】これらの報告されている方法を用いて自動または半自動でビデオの特徴をインデクスとしてビデオに記述、付加する。

【0024】ビデオ特徴(インデクス)情報が記述、付加されたビデオは、階層的構造化処理部103において、階層的に構造化する。ビデオ特徴(インデクス)情報を用いない単純な階層構造の例を図2に示す。この例は、ビデオの時間解像度を用いたものを示している。上層に向かうほど時間解像度が粗くなり(時間サンプリング間隔が大きくなる)、例えば、最上位層では、最下位層の1000倍速に相当する。

【0025】また、ビデオ特徴（インデクス）情報を用いることにより、単純な時間サンプリングから代表画面によるサンプリングが可能となる。上位層は下層の代表画面で構成される。

【0026】次に、図3に示した階層は、ビデオ特徴（インデクス）の物理的な情報を用いた階層を示している。輝度、色の変化、被写体の動き等の物理的な情報を用い特徴の連続性から順次最上位層の映像を時間的に分割した例であり、最下位層では、フレーム、画素等のレベルまで分割し、階層的に構造化しているものである。

【0027】また、ビデオには第4図に示したような下層の物理的な層から上層の意味的な層をなす階層構造が存在すると考えられる。この構造は、各層における情報が図に示した単位（画素、セグメント、フレーム、ショット、シーン、ストーリー）で管理されていると考えられる。

【0028】この場合の例を映画を用いて具体的に説明すると、最上位層では「タイトル」が次に「代表的なシーン」がそして各「ショット」という具合にビデオは意味的に階層的に構造化する。

【0029】以上説明した階層構造は、構造化条件情報蓄積部104に蓄積されている構造化条件に基づいて階層的構造化処理部103において構成される。階層的構造化処理部103において構造化されたビデオは、構造化ビデオ情報蓄積・管理部105に蓄積され、構造化情報可視化処理部106の可視化のためのビデオ情報源となる。

【0030】構造化情報可視化処理部106では、構造化ビデオ情報（階層構造）に基づいてビデオの可視化を行う。この可視化は可視化するための装置に依存するところが大きい。

【0031】以下、ビデオの構造を用いたビデオ情報の可視化についての実施例を説明する。ここで、可視化とは、検索者の興味に応じて可視化されるビデオ情報の形の変形を伴うものである。検索者の興味は、ビデオ情報アクセス手順によって構造化情報可視化手順へと入力される。この入力デバイスとしてマウス、ジョイスティック、キーボードからデータグローブまで様々対象となる。

【0032】これら入力デバイスからの信号を検索者からの興味の意志として表現する方法には、ビデオ情報に近づくことを表現する方法が考えられる。すなわち、ビデオ情報へズームする方法かまたは、ビデオ情報へ距離的に近づく方法である。

【0033】ズームする方法と距離的に近づく方法とは、「ビデオ情報へ近づくこと」が「ビデオ情報の詳細な部分へ近づくこと」であることと同じであるが、距離的に近づく方法は、遠近法の効果によりビデオ情報との距離を検索者に意識させることが可能な点で異なる。

【0034】可視化するためのシステムの容易性、安価

性からズームする方法がより一般的であるが、最近の計算機能力の向上から距離的に近づく方法の実現も可能である。距離的に近づく方法は、検索者とビデオ情報との位置関係が直感的に意識できること等の有効性から、以下この距離的に近づく方法を用いた例を示す。検索者の興味の意識の表現として距離メタファーを用いる。距離メタファーとは計算機上の仮想空間の距離の概念を導入に空間内における人と情報との関係を距離で比喩的に用いることである。

【0035】ビデオ検索のための仮想空間に検索者と可視化されたビデオ情報が存在する。可視化はビデオの構造情報を用いて行われている。遠くからビデオ情報を検索した場合、検索者は空間上に存在するビデオのタイトルを観察することとなりこの単位で検索することとなる。興味を示したタイトルと他のタイトルとの関係（空間的位置関係）を意識しつつ、興味を示したタイトルへ近づくと、近づくにつれてストーリーを構成するシーンが観察可能となる。この距離ではシーン単位での検索である。

【0036】次に、より近づくことによりショット、フレーム、セグメント、画素等の細かさでビデオ映像が観察でき、検索可能である。ビデオ情報として映画を例に取り、具体的に説明する。映画は、ビデオ情報解析手順により解析され、ビデオ構造化手順によって図4に示したような階層構造をもつ構造に構造化されている。

【0037】図5に可視化の様子を示す。図5において、501は映像の一つのショットを時空間画像として表している。一つの時空間画像501の拡大図を502に示す。一般の映画には、千数百個のショットが存在する。これらのショットを時間順に並べたもの（時間順配列ショット群）が503である。並べ方はいろいろ考えられるが、この例の時間順配列ショット群503では向かって左上から右下へ順次四角形を形成するように並べたものである。

【0038】さて、時間順配列ショット群503を検索者が仮想空間内の遠方から観察した場合、時間順配列ショット群503は、観察画像506として観察できる。すなわち、この場合、映画のタイトルを観察することになる。検索者は遠方より映画のタイトルを観察し、所望（興味ある）の映画を認識し、より多くの情報を得るために近づく行為を行う。観察画像506に近づくと、観察画像506は505のように観察できる。505は映画の代表的なシーン（504）から形成されている。この段階で検索者は、所望の（興味ある）の映画のシーンを認識することが可能である。

【0039】同様に、より多くのシーンの情報を得るために505へ近づくと、時間順配列ショット群503が観察可能となり、ショット単位で観察または検索可能となる。これら行為は、映画情報までの距離による関係から得られる情報を変化させているため観察または検索過

程が連続しており、観察者は、映画との距離（情報の多さ）を意識しつつ検索可能となる。

【0040】この視覚化は、検索による空間移動の行為をより自然に検索者に感じさせるために、空間の遠近法による表現と、検索者の視覚特性（視距離に応じて空間分解が異なる）を用いている。タイトルはシーンを用いて表現し、シーンはショットで表現することで実現している。表現は、シーン、ショットの画像の輝度値、色等を適宜変換することによって行うことが可能である。

【0041】以上までは、画像の可視化について説明してきたが、音についても前記した実施例に含まれる。実施例での音の扱いは、遠くから映画を観察すれば、タイトル音（タイトル音楽）が近づくにつれてシーンを代表する音が、ショットの音がそれぞれ距離移動に対して連続的に聞こえる。

【0042】以上、本発明を実施例に基づき具体的に説明したが、本発明は、前記実施例に限定されるものではなく、その要旨を逸脱しない範囲において種々変更可能であることは言うまでもない。

【0043】

【発明の効果】本願において開示される発明のうち代表的なものによる効果を簡単に説明すると、以下のとおりである。

【0044】ビデオ情報源のビデオ特徴を解析し、該解析結果を階層的に構造化し、該ビデオ情報源の階層構造を可視化し、該可視化ビデオ情報へアクセスするビデオ情報検索方法において、前記ビデオ情報解析処理は、ビデオの画像、音声、字幕、タイムコード、外部入力信号\*

＊を解析し、ビデオ情報の空間的变化特徴及び時間的变化特徴を抽出し、前記ビデオ構造化処理は、ビデオの画像、音声、字幕、タイムコード、外部入力信号を解析し、ビデオ情報の空間的变化特徴及び時間的变化特徴を抽出した条件による構造条件を基に階層的に構造化することより、ビデオ情報から所望のビデオまたはシーンを検索する場合に、検索者と所望情報との位置関係を検索者に意識させつつビデオ構造の積極的利用を行うので、検索過程を直感的にしかも効率よく行うことができる。

10 【図面の簡単な説明】

【図1】本発明にかかる情報処理装置（コンピュータ）を用いてビデオ情報検索方法を行う一実施例の機能構成及びその処理手順を示すブロック機能構成図である。

【図2】本実施例のビデオ特徴（インデクス）情報を用いない単純な階層構造の例を示す図である。

【図3】本実施例のビデオ特徴（インデクス）の物理的な情報を用いた階層を示す図である。

【図4】本実施例のビデオ情報の意味的階層構造を示す図である。

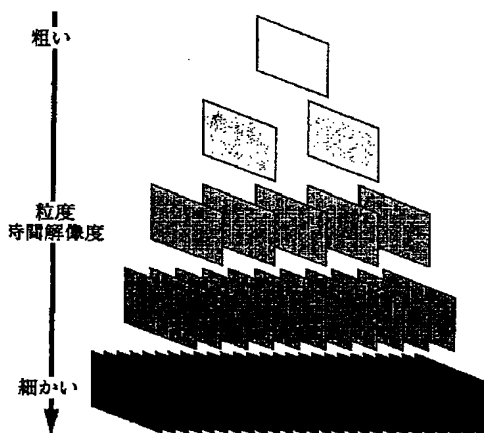
20 【図5】本実施例のビデオ情報の構造化情報を用いた可視化（映像インターフェイス）を説明するための図である。

【符号の説明】

101…ビデオ特徴判断部、102…ビデオ情報解析処理部、103…階層的構造化処理部、104…構造化条件情報蓄積部、105…構造化ビデオ情報蓄積・管理部、106…構造化情報可視化処理部、107…ビデオ情報アクセス部。

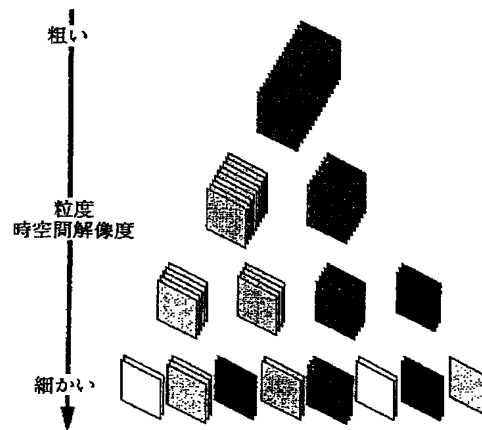
【図2】

図2



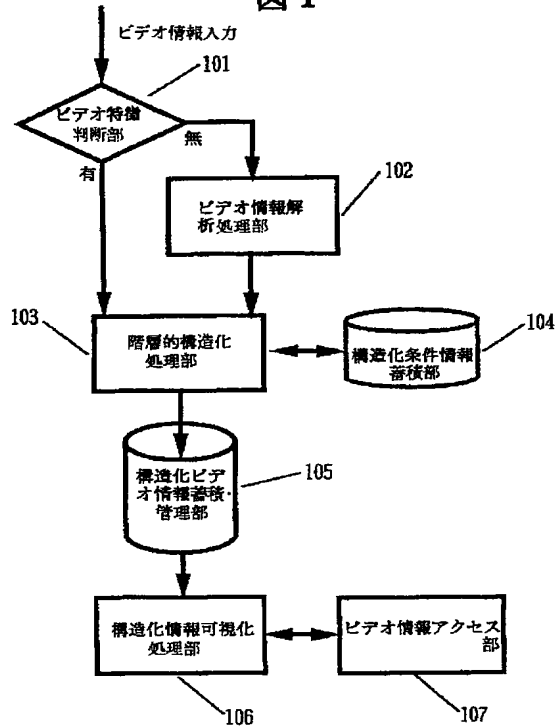
【図3】

図3



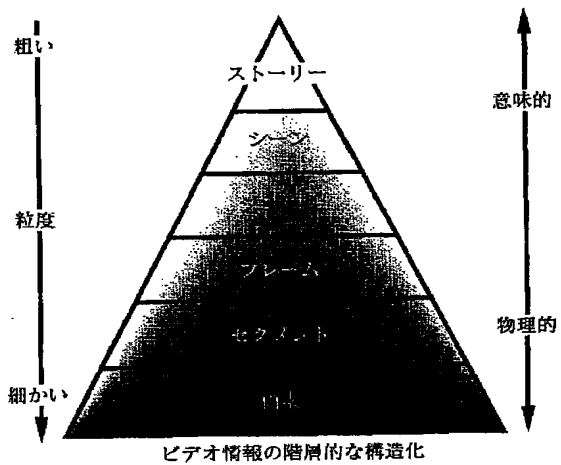
【図 1】

図 1



【図 4】

図 4



【図 5】

図 5

